DOCUMENT RESUME

ED 421 531 TM 028 857

AUTHOR	Bay, Luz; Loomis, Susan Cooper; Wang, Tianyou
TITLE	Scoring with Nominal Missing-Response Parameters: Its Effects on Achievement Levels Set on the National Assessment of Educational Progress (NAEP).
PUB DATE	1995-04-22
NOTE	43p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Comparative Analysis; *Cutting Scores; High School Seniors; High Schools; *Item Response Theory; National Competency Tests; Probability; *Scoring; Test Results
IDENTIFIERS	Missing Data; *National Assessment of Educational Progress; *Nominal Response Model; Three Parameter Model; Two Parameter Model

ABSTRACT

This study examines the effects on the National Assessment of Educational Progress (NAEP) achievement levels of using item response theory (IRT) models that have nominal missing-response parameters. It compares cutpoints based on item parameters that were fitted using two different models. The first set of cutpoints were based on parameters for the two- and three-parameter logistic model, and the second set of cutpoints were based on R. Bock's (1972) nominal model. Data are from the 1992 NAEP in mathematics and reading for grade 12. For reading, data included 1,966 responses to a block of items, and for mathematics, data included 2,192 responses. Other data were the item-by-item ratings by each panelist who participated in the Achievement Levels Setting process for NAEP 1992. For each subject, the cutpoints set using the probability curves obtained by the different models were compared. The percent of students scoring at each level were also compared. For reading, the logistic model, when fitted to the data, converged in 25 iterations and yielded a marginal reliability of 0.67 with maximum information of 4.3 at theta equals -0.5. The nominal model converged in 88 iterations and yielded a marginal reliability of 0.85 with a maximum information of 14.2 at theta equals 0. For mathematics, the logistic model converged in 184 iterations and yielded a marginal reliability of 0.61 with maximum information of 23.9 at theta equals 1.5. The nominal model, with mathematics data, converged in 46 iterations, and yielded a marginal reliability of 0.62, with maximum information of 7.3 at theta equals -0.2. When the percentages of students scoring at or above each cutpoint were compared, none scored at the Advanced level using the nominal model. These preliminary results suggest the direction of future studies, but cannot be generalized to the NAEP assessment program. An appendix contains achievement levels descriptions. (Contains three tables, six figures, and five references.) (SLD)

ED 421 531

Scoring with Nominal Missing-Response Parameters: Its Effects on Achievement Levels Set on the National Assessment of Educational Progress (NAEP)

Luz Bay

Susan Cooper Loomis

Tianyou Wang

American College Testing

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Luz Bay
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at the annual meeting of the American Educational Research Association

22 April 1995

TM028857

Scoring with Nominal Missing-Response Parameters: Its Effects on Achievement Levels Set on the National Assessment of Educational Progress (NAEP)

Luz Bay
Susan Cooper Loomis
Tianyou Wang

American College Testing

Background of the Study: The National Assessment of Educational Progress (NAEP) is one of the large-scale assessment programs that develop tests that combine multiple-choice and open-ended items. In the 1992 NAEP in Mathematics, for example, there were 99 multiple-choice, 54 short answer, and five extended answer items for grade four, 118 multiple-choice, 59 short answer, and six extended answer items for grade eight, and 115 multiple-choice, 58 short answer, and six extended answer items for grade twelve. Overall, 64% of the items were multiple-choice, 33% were short answer, and three percent were extended answer. The NAEP items were administered in blocks, each of which is a combination of multiple-choice and constructed response items.

In 1994, Swinton, in his paper titled *Scoring with Nominal Missing-Response Parameters*, reported that the increasing proportion of open-ended items in the NAEP corresponded to a rise in the number of nonresponse to those items. He further stated that "this problem is exacerbated when multiple-choice and open-ended items are presented in the same block, with at least one multiple-choice item following an open-ended item" (p. 1). This "exacerbation" results from

the potential for an examinee to attempt the multiple-choice items first, then go back to open-ended items (if there is more time). This invalidates NAEP's traditional scoring approach that considers omitted items to be incorrect, and "Not Reached" items to be missing. Items treated as missing do not affect the level of an examinee's ability estimate. Such a scoring approach is only valid under the assumption that students attempt items in sequential order. Swinton (1994) suggested that if one cannot reliably distinguish omitting- from nonreaching-behavior, an option is to attempt to model nonresponse in the scoring process.

The study reported in this paper is an extension of Swinton's 1994 study. It examines the effects on the NAEP achievement levels of using IRT models that have nominal missing-response parameters. It compares cutpoints based on item parameters that were fitted using two different models. The first set of cutpoints were based on parameters for the two- and three-parameter logistic model. The second set of cutpoints were based on Bock's (1972) nominal model.

Data[1]: Data used for this study are from the 1992 NAEP in Mathematics and Reading for grade 12. For the purposes of this study, only one block of items from each subject was used. For Reading, the data included the responses of 1,966 grade-12/age-17 students to items in block RD. This block included three multiple-choice (M) items, five short open-ended (O) items, and one extended

[1]The authors wish to thank the Center for Assessment of Educational Progress and the Educational Testing Service for the data sets that they provided for this study.

response (E) item. For Mathematics, the data set included responses of 2,192 grade-12/age-17 students to items in block M15. This block included six multiple-choice items, three short open-ended items, and one extended response item. Codes for responses to multiple-choice items correspond to the five choices, plus three categories corresponding to omitted items, not reached items, and multiple response items. Codes for responses to open-ended items correspond to the score levels, plus three categories corresponding to omits, not reached, and off-task. The order of the different types of items in each block and their nonresponse rates (*i.e.*, rates of omits and not reached) are in Table 1. Notice that the nonresponse rate is very high in Reading, especially for those items that come later in the block.

The other data sets that were used are the item-by-item ratings by each panelist who participated in the Achievement Levels-Setting (ALS) process in NAEP Mathematics and Reading in 1992. There were 10 raters in grade 12 Reading and 11 raters in grade 12 Mathematics. The data for each rater included a modified Angoff rating for each dichotomous item.[2] At each achievement level, Basic, Proficient, or Advanced, the rater provided his/her best estimate of the probability that a student performing at the lower borderline of each level would respond to the item correctly. Although this was done in three rounds, only the third round ratings were used in setting the cutpoints. Thus, in this study, only the third round ratings were used.

---

[2]Polytomous items were rated using the paper selection method. However, polytomous items will not be considered for this study for reasons that will be discussed in the Method section.

Method: The responses to dichotomous items were recoded so that there were only four response codes: "1"=correct, "0"= incorrect, "n"=not reached, and "o"=omit. Two different IRT models were fit to each response data set. Thissen's (1990) MULTILOG PC program was used to estimate the parameters and score the test. After the parameters were obtained, the probability functions were used to map the modified Angoff ratings to compute the cutpoints corresponding to the Achievement Levels Descriptions of Basic, Proficient, and Advanced. (Please see Appendix A for the Achievement Levels Descriptions.)

The first model that was fitted was the two-parameter logistic model, expanded to three parameters for multiple-choice items. In this model, omits were considered wrong and not reached were coded missing. This approximates the current NAEP scoring practice. Figure 1 shows a typical item characteristic curve (ICC) based on a three parameter logistic model. The second model that was fitted was Bock's (1972) nominal model. Unlike the previous model, Bock's model does not yield a logistic trace line for each item. Instead, it produces curves that are ratios of a category-specific exponential to the sum of the exponentials for each category. Three categories were used for this model: 1 = no response, 2 = incorrect, and 3 = correct. The "no response" category is the combination of omits and not reached. A typical set of response characteristic curves for the nominal model is shown in Figure 2. Each of these models were fitted to each of the response data sets. In each case a normal prior was assumed.

The graded response (GR) model (Samejima, 1969) was considered for this

study. In the GR model an *a priori* ordering of the categories is required, but there is no completely unambiguous way of ordering nonresponse and incorrect categories. For this reason, the GR model was not included for comparison.

Using the parameters estimated for each model, the probabilities of a correct response at each point on the θ-scale were summed to produce an expected test score. Thus, for the logistic model, the ICCs were added together to form a test score function (TSF). For the nominal model, the probability curves for the correct responses were summed to form a TSF. The TSFs were used to map the modified Angoff ratings to the theta scale to produce the cutpoints.

To map the modified Angoff ratings to the θ-scale, let

I = the number of items, and

J = the number of raters.

Suppose

r_{Xij} = raters j's estimate that a student performing at the borderline of the X achievement level will respond to item i correctly;

where X = Basic, Proficient, or Advanced,

$i = 1, 2, ..., I$,

$j = 1, 2, ..., J$.

These estimates or ratings are summed across items, and the sums are averaged across raters. If

$$r_X = \frac{\sum_{j=1}^{J} \sum_{i=1}^{I} r_{X_{ij}}}{J}$$

then, r_X is mapped to the theta scale using the TSF. The value θ_X is the lower borderline, or cutpoint, of the X achievement level. For example, using the TSF in Figure 3, if $r_{Proficient} = 4.36$ then the lower borderline of the Proficient achievement level is $\theta_{Proficient} = -0.40$.

Polytomous items were not used for this study because they cannot be included in the TSF for the nominal model. The current procedures use the partial credit (PC) model for polytomous items. The PC model produces a probability curve for each score level of a polytomous item. Suppose a polytomous item has four score levels and, given θ, the probability of getting a score of n, where n = 1, 2, 3, or 4, is $P(X=n|\theta)$ based on the partial credit model. Then sum

$$\sum_{n=1}^{4} nP(X=n|\theta)$$

will form the expected score curve for that item. At each value of θ, the expected score for polytomous items and the probability of getting correct answers for dichotomous items are summed to produce the TSF. When using the nominal model with a nonresponse category there is a problem in forming the expected

score for a polytomous item.[3] There is one polytomous item in Reading and one polytomous item in Mathematics. Thus, eight items in Reading and 10 items in Mathematics were used for the analyses.

For each subject, the cutpoints set using the probability curves obtained the different models were compared. The percent of students scoring at each level were also compared.

Results: When a logistic model with 19 total, free parameters (three parameters for each multiple-choice item, and two for each short-answer open-ended item) was fitted to the Reading data, it converged in 25 iterations and yielded a marginal reliability of .67 with maximum information of 4.3 at $\theta = -.5$. There are three multiple-choice items in Reading. The pseudo-chance level (*i.e.*, *c*) parameter for item 2 was estimated to be .29, and 0 for both items 4 and 7. When the nominal model, with 32 free parameters (four for each item), was fitted to the Reading data, it converged in 88 iterations, and yielded a marginal reliability of .85. It has a maximum information of 14.2 at $\theta = 0$. The estimated parameters for both models are listed on Table 2R.

For Mathematics, a logistic model, with 28 free parameters, converged in 184 iterations and yielded a marginal reliability of .61 with maximum information

---

[3]For a more detailed description of computing the cutpoints, please see Chapter 3 of the *Setting Achievement Levels of on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing: A Technical Report on Reliability and Validity*

of 23.9 at θ = 1.5. All but one of the items fitted with the three-parameter logistic model had nonzero *c* parameters. Item 4 was fitted with 3PL because of the nature of the item even though it is an open-ended item; however, the *c* parameter was estimated to be 0. When the nominal model, with 40 free parameters, was fitted it converged in 46 iterations and yielded a marginal reliability of .62, with maximum information of 7.3 at θ = -.2. The estimated parameters for Mathematics are given in Table 2M.

An *a priori* ordering of categories was not required for the nominal model. *A posteriori*, the nonresponse category scaled lower than the incorrect category in each item in both subjects. This is consistent with Swinton's (1994) results.

For each subject, the item-by-item ratings were summed across items and averaged across raters at each achievement level. These numbers are listed in column three of Table 3. Using the TSF obtained from each model, the ratings were mapped to the θ scale. The cutpoints on the θ scale are in column four for the logistic model, and column five for the nominal model in Table 3. Notice that in each case (except in Mathematics at the Basic level) the cutpoints were higher when the nominal model was used.

The TSFs using the different models are shown in Figures 4R and 4M. Notice that in each case, the expected score based on the logistic model was generally higher than the expected score based on the nominal model especially at the upper part of the θ-scale. In this study, the ICC was usually higher than the correct response curve from the nominal model. A typical example is shown in

Figure 5. It was observed, however, that for items with very high nonresponse rates the correct response curves converged together at high values of θ. An example of this is item number nine in Reading as seen in Figure 6.

Instead of comparing the cutpoints *per se*, a common way of comparing cutpoints is by comparing the percents of students scoring at-or-above each cutpoint. This information is provided in columns five and seven of Table 3. These percentages are of students in the data set who scored at or above each cutpoint. In both Reading and Mathematics none of the students scored at the Advanced level using the nominal model. There were more students scoring at or above the cutpoints set using the logistic model, with the exception of the Basic level in Reading. The differences in the percentages were very small at the Basic level for each subject. The values were 1.48% (29 of 1,966 students) in Reading and 2.15% (47 of 2,192 students) in Mathematics. The largest difference in Reading was at the Proficient level (32.71%), and the smallest was at the Basic level. The differences in percents at or above were not as large in Mathematics as they were in Reading. The largest difference in Mathematics was also at the proficient level, but it was only 5.06%.

Significance of the Study: A number of studies have been done regarding student test-taking behavior, and even more have been done on different item types. In 1993, Swinton discussed students' test-taking strategies when there is a combination of different item types in a test. This inquiry is a study in progress,

and it is an attempt to explore the effects on achievement levels of using IRT models that include nonresponse parameters when one cannot reliably distinguish "omitting" from "non-reaching" behavior due to students' test taking strategies on multiple item types.

The results presented here are preliminary. Moreover, since the IRT scaling using logistic model only approximates the current NAEP scaling, the results cannot be generalized to the NAEP assessment program.

References:

Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika, 37*, 29-51.

Samejima, F. (1969). *Estimation of latent ability using response pattern of graded scores* (Psychometric Monograph No. 17). Iowa City, IA: Psychometric Society.

Swinton, S. R. (April, 1994). *Scoring with nominal missing-response parameters.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Swinton, S. R. (April, 1993). *Nonresponse to NAEP items of different formats as a function of ability, age, gender, and ethnicity.* Paper presented at the annual meeting of the American Educational Research Association.

Thissen, D. M. (1991). *MULTILOG: Multiple categorical item analysis and test scoring using item response theory* (Version 6.0) [Computer program]. Mooresville, IN: Scientific Software.

Table 1: Order of Different Types of Items and Their Nonresponse Rates

Item No.	Mathematics		Reading	
	Item Type	% Nonresponse	Item Type	% Nonresponse
1	O	1	O	7
2	M	1	M	2
3	O	1	O	18
4	O	2	M	8
5	M	1	O	14
6	M	1	E	38
7	M	4	M	22
8	M	3	O	39
9	M	2	O	44
10	M	4		
11	E	17		

Table 2R: Estimated Item Parameters for Reading

Item No.	Model								
	Logistic			Nominal					
	a	b	c	a_1	a_2	a_3	c_1	c_2	c_3
1	1.67	-0.52		-0.71	0.25	0.46	-1.39	0.32	1.08
2	1.40	-0.61	0.29	-1.29	0.49	0.79	-2.73	0.73	1.99
3	0.90	-1.11		-0.92	0.35	0.57	-0.53	-0.45	0.99
4	0.59	-0.67	0.00	-3.12	1.47	1.64	-3.43	1.43	1.99
5	1.46	-1.15		-3.36	1.58	1.77	-2.19	0.35	1.85
7	0.67	-0.19	0.00	-2.12	0.91	1.22	-0.89	0.41	0.48
8	0.76	0.69		-3.68	1.14	2.54	0.26	0.61	-0.87
9	1.58	0.06		-4.06	0.21	3.85	1.23	0.71	-1.93

Table 2M: Estimated Item Parameters for Mathematics

Item No.	Model								
	Logistic			Nominal					
	a	b	c	a_1	a_2	a_3	c_1	c_2	c_3
1	0.75	-1.45		-0.64	0.00	0.63	-2.65	0.77	1.88
2	1.05	-0.25	0.21	-1.21	0.10	1.10	-3.45	1.30	2.14
3	1.04	0.26		-1.14	0.12	1.01	-3.03	1.63	1.39
4	0.41	0.11	0.00	-1.69	0.53	1.15	-3.74	1.90	1.84
5	1.98	1.68	0.19	-1.83	0.68	1.15	-4.50	2.84	1.66
6	0.52	0.08	0.00	-2.95	1.11	1.84	-6.70	3.40	3.30
7	0.60	2.32	0.10	-1.03	0.24	0.80	-2.10	1.76	0.33
8	2.07	1.51	0.10	-1.66	0.31	1.35	-2.74	2.27	.48
9	17.25	1.57	0.28	-1.51	0.67	0.85	-3.53	2.19	1.35
10	1.01	0.93	0.02	-1.79	0.08	1.70	-2.58	2.00	0.57

Table 3: Cutpoints

Subject	Level	Ratings	Model			
			Logistic		Nominal	
			Cutpoint	%≥	Cutpoint	%≥
Reading	Basic	2.85	-0.94	87.59	-0.83	89.07
	Proficient	5.15	-0.11	58.5	0.69	25.79
	Advanced	6.84	0.77	13.89	2.89	0.00
Mathematics	Basic	3.71	-0.09	55.89	-0.14	53.74
	Proficient	6.46	1.38	7.39	1.63	2.33
	Advanced	8.56	2.00	0.55	3.80	0.00

Figure 1
A Typical Item Characteristic Curve from a 3-Parameter Logistic Model

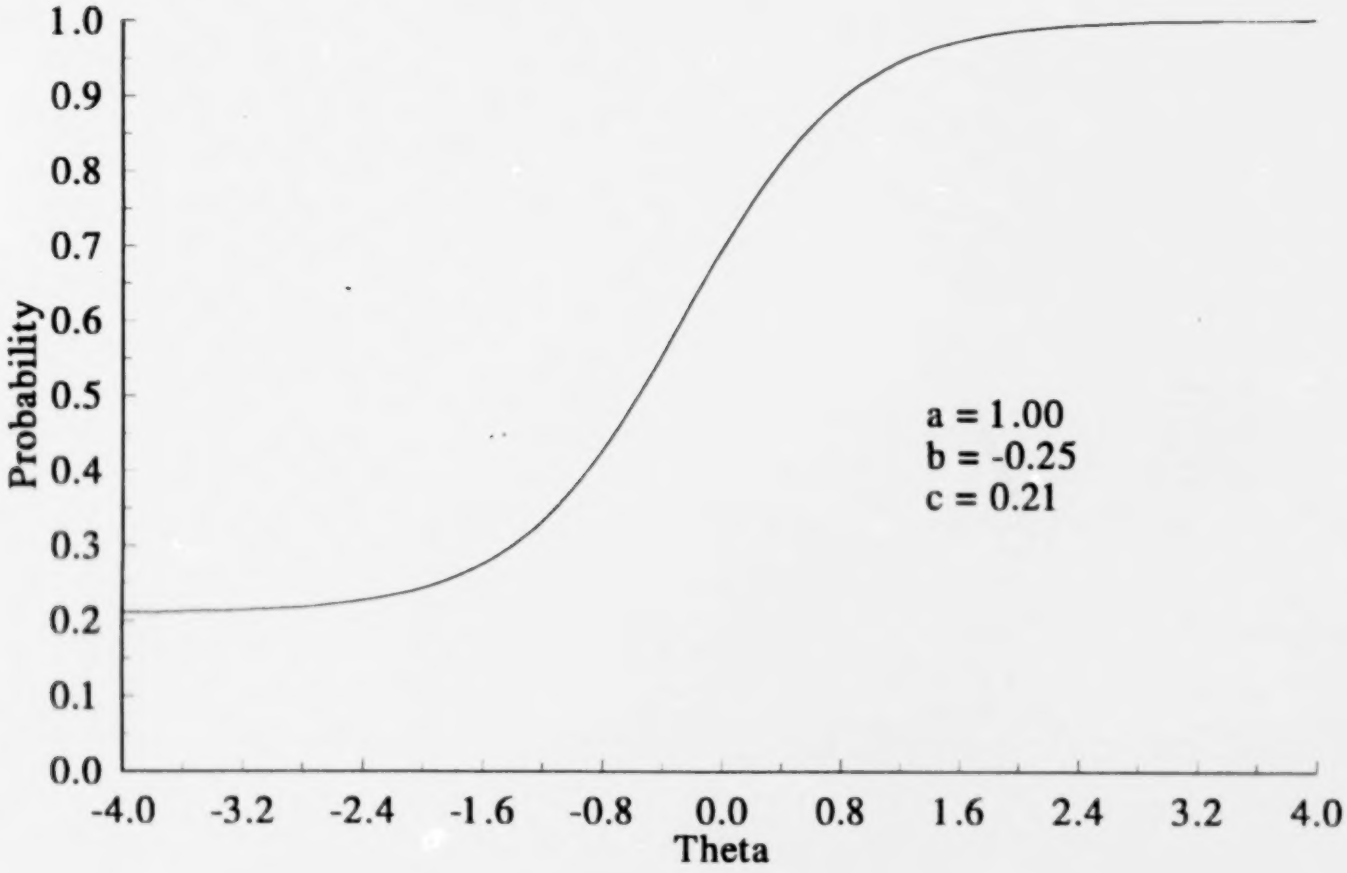


Figure 2
A Typical Set of Response Curves from a Nominal Model w/ Three Categories

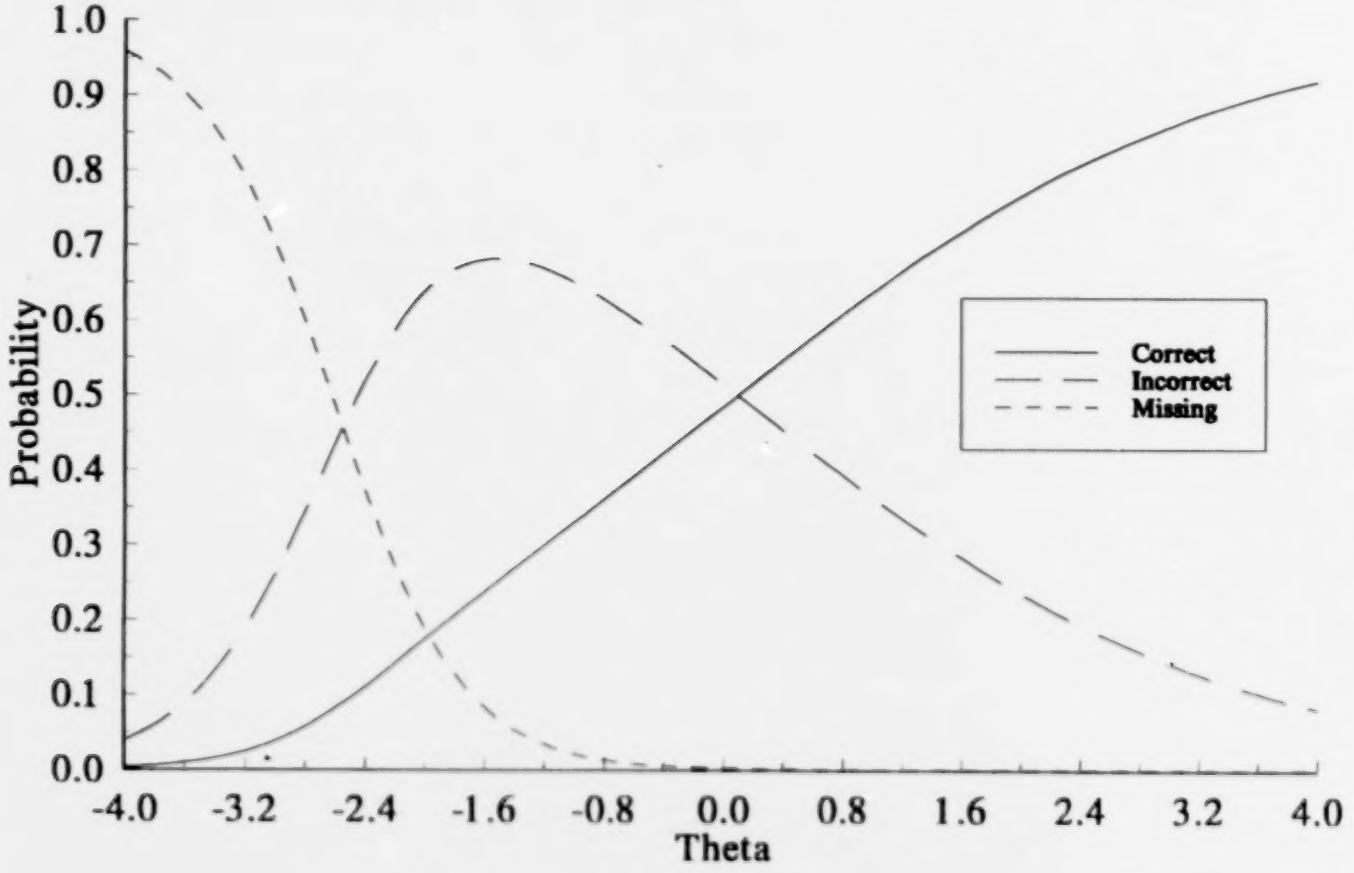


Figure 3
Mapping the Ratings to the Theta Scale Using the Test Score Function (TSF)

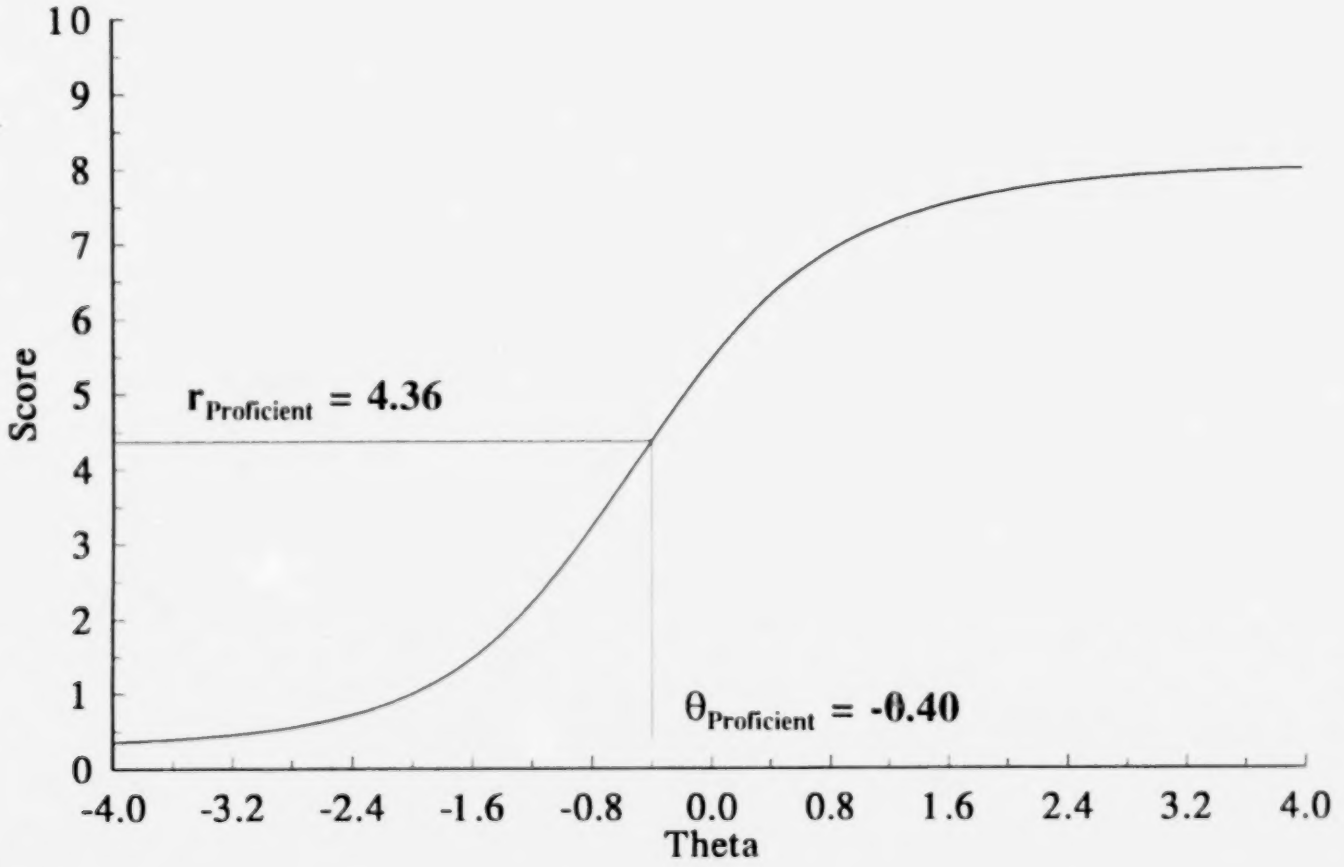


Figure 4R
Test Score Functions (TSFs) for Reading

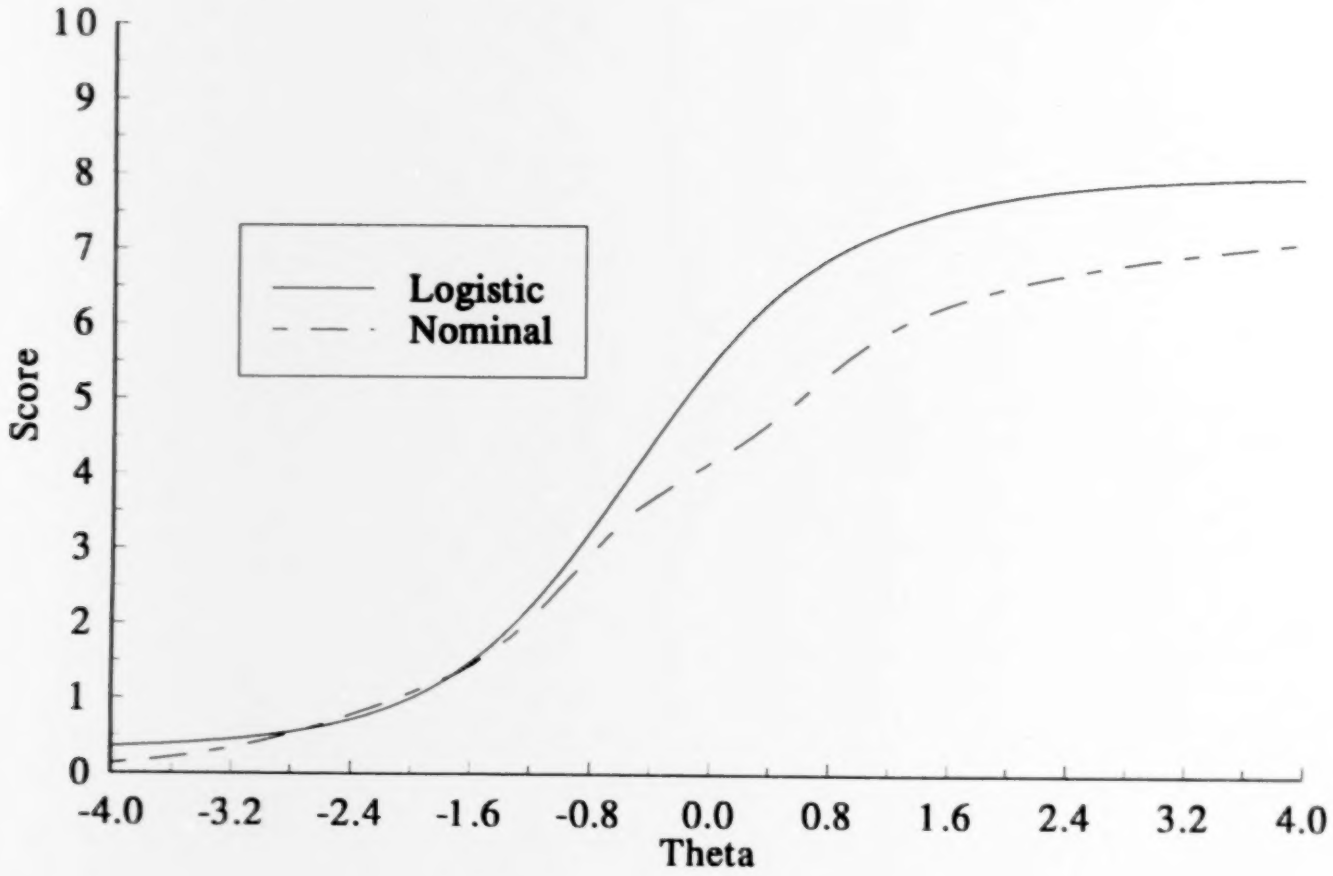


Figure 4M
Test Score Functions (TSFs) for Mathematics

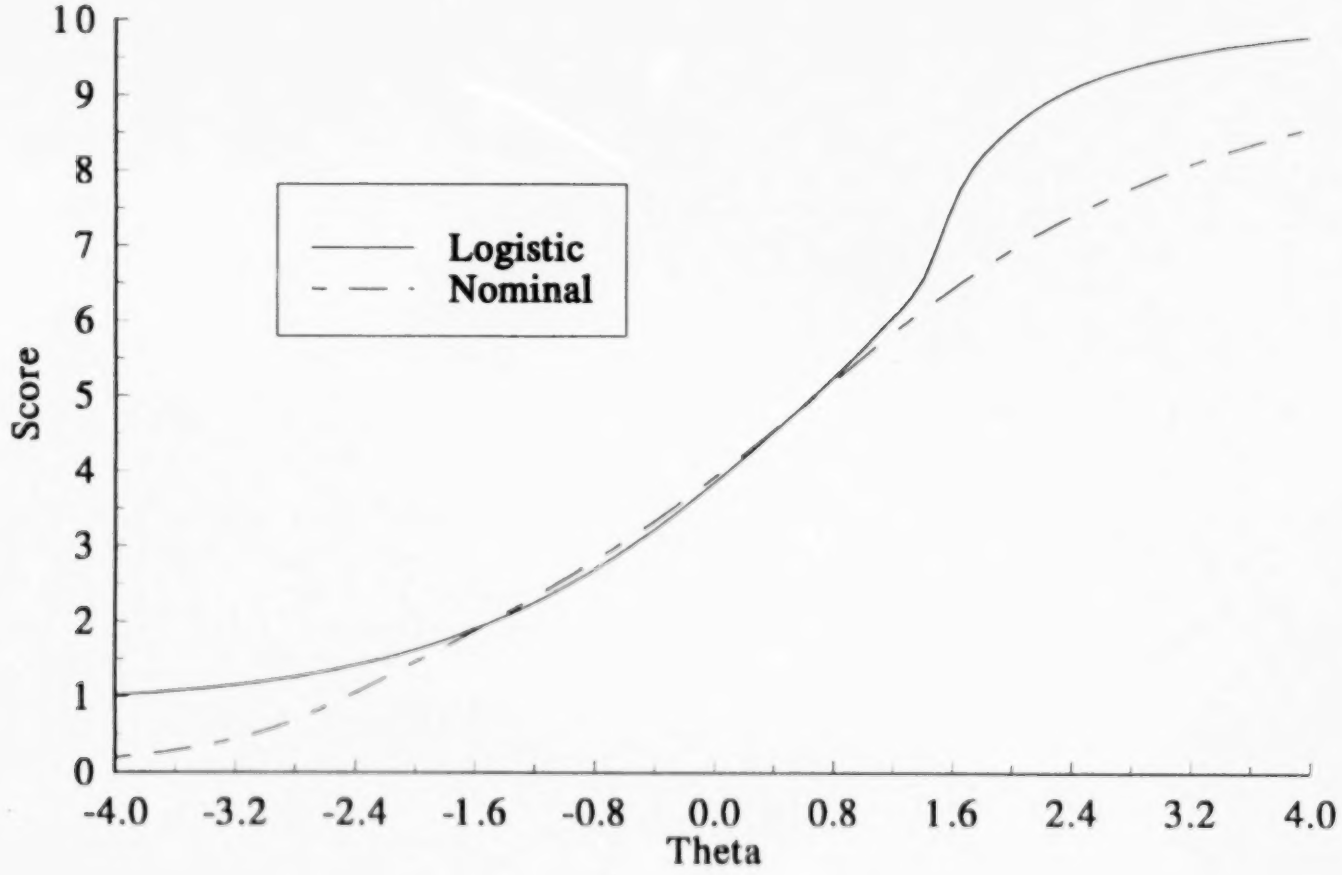


Figure 5
Reading Item Number 7

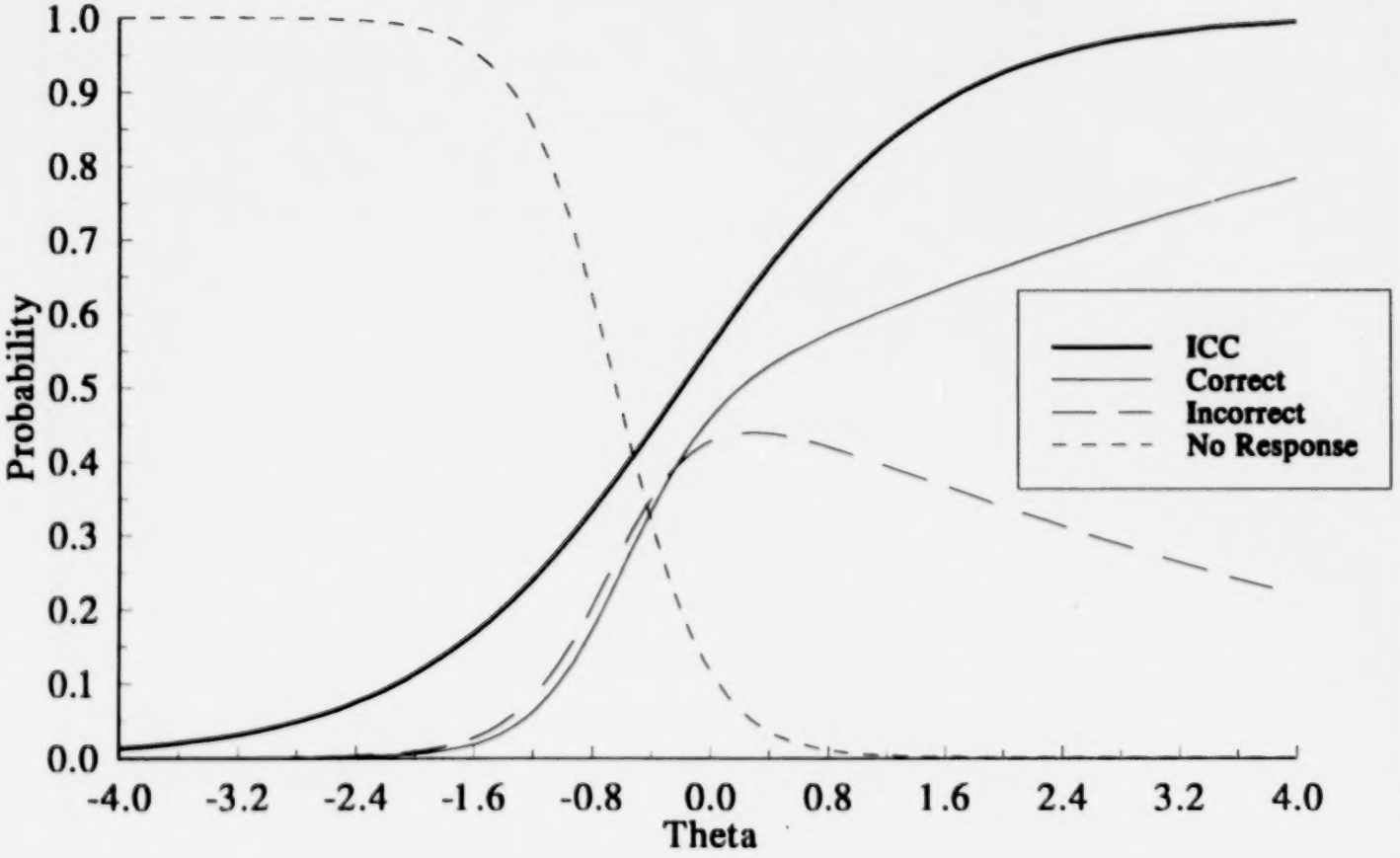
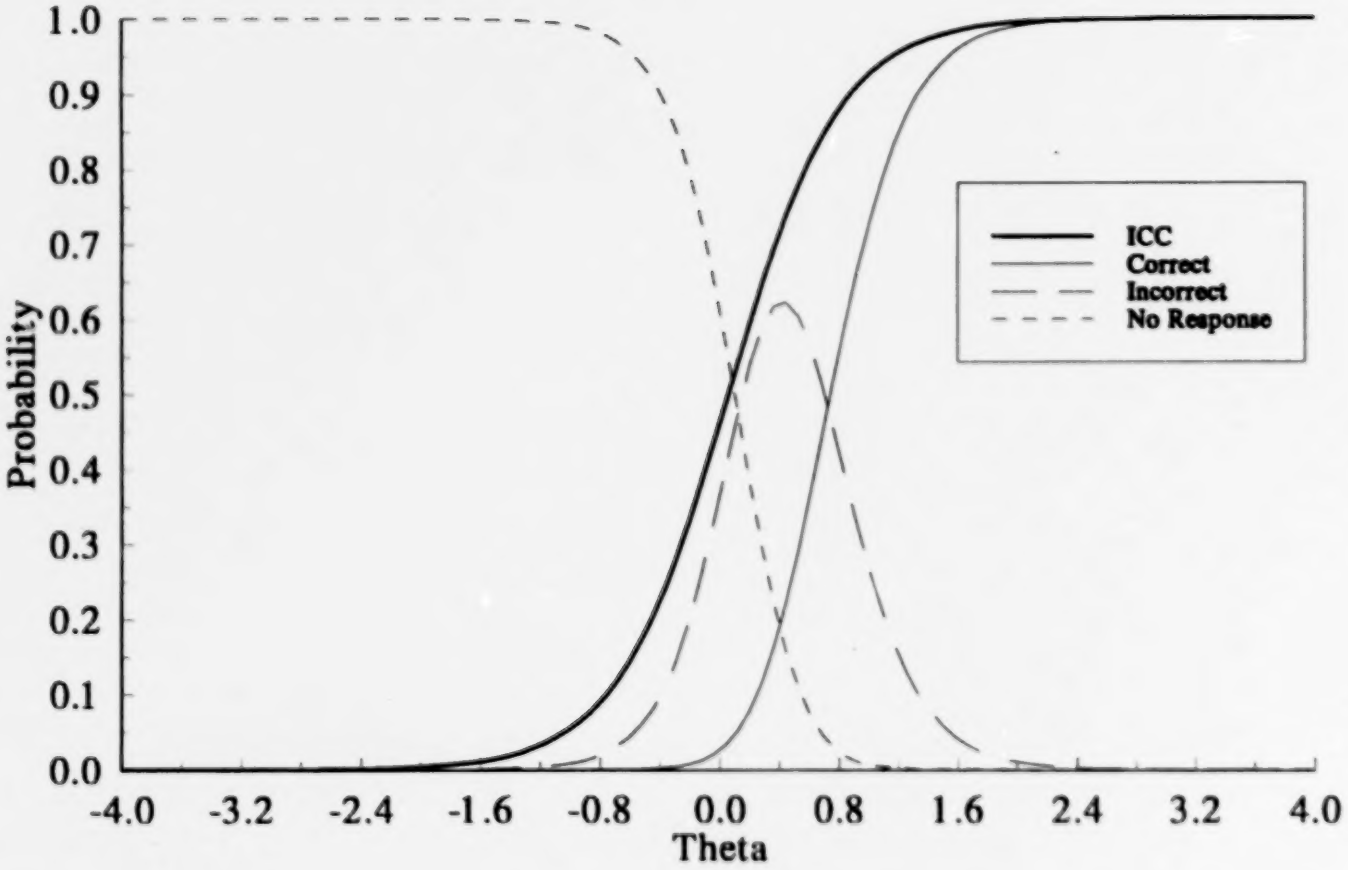


Figure 6
Reading Item Number 9



Appendix A

Achievement Levels Descriptions

Policy Definitions of Achievement Levels

Proficient.

This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real world situations, and analytical skills appropriate to the subject matter.

Basic.

This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

Advanced.

This level signifies superior performance beyond proficient.

Reading

Descriptions of Reading Achievement Levels for Basic, Proficient, and Advanced Student Performance on the 1992 NAEP

Preamble

Reading for meaning involves a dynamic, complex interaction between and among the reader, the text, and the context. Readers, for example, bring to the process their prior knowledge about the topic, their reasons for reading it, their individual reading skills and strategies, and their understanding of differences in text structures.

The texts used in the reading assessment are representative of common real world reading demands. Students at Grade 4 are asked to respond to literary and informational texts which differ in structure, organization, and features. Literary texts include short stories, poems, and plays that engage the reader in a variety of ways, not the least of which is reading for fun. Informational texts include selections from textbooks, magazines, encyclopedias, and other written sources whose purpose is to increase the reader's knowledge.

In addition to literary and informational texts, students at Grades 8 and 12 are asked to respond to practical texts (e.g., bus schedules or directions for building a model airplane) that describe how to perform a task.

The context of the reading situation includes the purposes for reading that the reader might use in building a meaning of the text. For example, in reading for literary experience, students may want to see how the author explores or uncovers experiences, or they may be looking for vicarious experience through the story's characters. On the other hand, the student's purpose in reading informational texts may be to learn about a topic (such as the Civil War or the oceans) or to accomplish a task (such as getting somewhere, completing a form, or building something).

The assessment asks students at all three grades to build, extend, and examine text meaning from four stances or orientations:

- **Initial Understanding**—Students are asked to provide the overall or general meaning of the selection. This includes summaries, main points, or themes.

- **Developing Interpretation**—Students are asked to extend the ideas in the text by making inferences and connections. This includes making connections between cause and effect, analyzing the motives of characters, and drawing conclusions.

- **Personal Response**—Students are asked to make explicit connections between the ideas in the text and their own background knowledge and experiences. This includes comparing story characters with themselves or people they know, for example, or indicating whether they found a passage useful or interesting.

- **Critical Stance**—Students are asked to consider how the author crafted a text. This includes identifying stylistic devices such as mood and tone.

These stances are not considered hierarchical or completely independent of each other. Rather, they provide a frame for generating questions and considering student performance at all levels. All students at all levels should be able to respond to reading selections from all of these orientations. What varies with students' developmental and achievement levels is the amount of prompting or support needed for response, the complexity of the texts to which they can respond, and the sophistication of their answers.

Introduction

The following achievement-level descriptions focus on the interaction of the reader, the text, and the context. They provide some specific examples of reading behaviors that should be familiar to most readers of this document. The specific examples are not inclusive; their purpose is to help clarify and differentiate what readers performing at each achievement level should be able to do. While a number of other reading achievement indicators exist at every level, space and efficiency preclude an exhaustive listing. It should also be noted that the achievement levels are cumulative from Basic to Proficient to Advanced. One level builds on the previous levels such that knowledge at the Proficient level presumes mastery of the Basic level, and knowledge at the Advanced level presumes mastery at both the Basic and Proficient.

Basic 212

Fourth-grade students performing at the **basic level should demonstrate an understanding of** the overall meaning of what they read. **When reading texts appropriate for 4th graders, they should be able to make relatively obvious connections between the text and their own** experiences. **For example, when reading literary text, they should be able to tell what the** story is generally about—providing details to support their understanding—and be able to connect aspects of the stories to their own experiences. **When reading informational text,** basic-level 4th graders should be able to tell what the selection is generally about or identify the purpose for reading it; provide details to support their understanding; and connect ideas from the text to their background knowledge and experiences.

Proficient 243

Fourth grade students performing at the **proficient level should be able to demonstrate an** overall understanding of the text, providing inferential as well as literal information. When reading text appropriate to 4th grade, **they should be able to extend the ideas in the text by** making inferences, drawing conclusions, and making connections to their own experiences. The connection between the text and what the student infers should be clear. For example, when reading **literary text, proficient-level 4th graders should be able to summarize the story,** draw conclusions about the characters or plot, and recognize relationships such as cause and effect. When reading **informational text, proficient-level students should be able to summarize** the information and identify the author's intent or purpose. **They should be able to draw** reasonable conclusions from the text, recognize relationships such as cause and effect or similarities and differences, and identify the meaning of the selection's key concepts.

Advanced 275

Fourth grade students performing at the **advanced level should be able to generalize about** topics in the reading selection and demonstrate an awareness of how authors compose and use literary devices. **When reading text appropriate to 4th grade, they should be able to judge** texts critically and, in general, give thorough answers that indicate careful thought. For example, when reading **literary text, advanced-level students should be able to make** generalizations about the point of the story and extend its meaning by integrating personal experiences and other readings with the ideas suggested by the text. **They should be able** to identify literary devices such as figurative language. **When reading informational text,** advanced-level 4th graders should be able to explain the author's intent by using supporting material from the text. **They should be able to make critical judgments of the form and** content of the text and explain their judgments clearly.

BEST COPY AVAILABLE

Basic 244

Eighth grade students performing at the basic level should demonstrate a literal understanding of what they read and be able to make some interpretations. When reading text appropriate to 8th grade, they should be able to identify specific aspects of the text that reflect the overall meaning, recognize and relate interpretations and connections among ideas in the text to personal experience, and draw conclusions based on the text. For example, when reading literary text, basic-level 8th graders should be able to identify themes and make inferences and logical predictions about aspects such as plot and characters. When reading informative text, they should be able to identify the main idea and the author's purpose. They should make inferences and draw conclusions supported by information in the text. They should recognize the relationships among the facts, ideas, events, and concepts of the text (e.g., cause and effect and chronological order). When reading practical text, they should be able to identify the main purpose and make predictions about the relatively obvious outcomes of procedures in the text.

Proficient 283

Eighth grade students performing at the proficient level should be able to show an overall understanding of the text, including inferential as well as literal information. When reading text appropriate to 8th grade, they should extend the ideas in the text by making clear inferences from it, by drawing conclusions, and by making connections to their own experiences—including other reading experiences. Proficient 8th graders should be able to identify some of the devices authors use in composing text. For example, when reading literary text, students at the proficient level should be able to give details and examples to support themes that they identify. They should be able to use implied as well as explicit information in articulating themes; to interpret the actions, behaviors, and motives of characters; and to identify the use of literary devices such as personification and foreshadowing. When reading informative text, they should be able to summarize the text using explicit and implied information and support conclusions with inferences based on the text. When reading practical text, proficient-level students should be able to describe its purpose and support their views with examples and details. They should be able to judge the importance of certain steps and procedures.

Advanced 328

Eighth grade students performing at the advanced level should be able to describe the more abstract themes and ideas of the overall text. When reading text appropriate to 8th grade, they should be able to analyze both meaning and form and support their analyses explicitly with examples from the text; they should be able to extend text information by relating it to their experiences and to world events. At this level, student responses should be thorough, thoughtful, and extensive. For example, when reading literary text, advanced-level 8th graders should be able to make complex, abstract summaries and theme statements. They should be able to describe the interactions of various literary elements (i.e., setting, plot, characters, and theme); to explain how the use of literary devices affects both the meaning of the text and their response to the author's style. They should be able critically to analyze and evaluate the composition of the text. When reading informative text, they should be able to analyze the author's purpose and point of view. They should be able to use cultural and historical background information to develop perspectives on the text and be able to apply text information to broad issues and world situations. When reading practical text, advanced-level students should be able to synthesize information that will guide their performance, apply text information to new situations, and critique the usefulness of the form and content.

BEST COPY AVAILABLE

Basic 269

Twelfth grade students performing at the basic level should be able to demonstrate an overall understanding and make some interpretations of the text. When reading text appropriate to 12th grade, they should be able to identify and relate aspects of the text to its overall meaning, recognize interpretations, make connections among and relate ideas in the text to their personal experiences, and draw conclusions. They should be able to identify elements of an author's style. For example, when reading literary text, 12th-grade students should be able to explain the theme, support their conclusions with information from the text, and make connections between aspects of the text and their own experiences. When reading informational text, basic-level 12th graders should be able to explain the main idea or purpose of a selection and use text information to support a conclusion or make a point. They should be able to make logical connections between the ideas in the text and their own background knowledge. When reading practical text, they should be able to explain its purpose and the significance of specific details or steps.

Proficient 304

Twelfth grade students performing at the proficient level should be able to show an overall understanding of the text which includes inferential as well as literal information. When reading text appropriate to 12th grade, they should be able to extend the ideas of the text by making inferences, drawing conclusions, and making connections to their own personal experiences and other readings. Connections between inferences and the text should be clear, even when implicit. These students should be able to analyze the author's use of literary devices. When reading literary text, proficient-level 12th graders should be able to integrate their personal experiences with ideas in the text to draw and support conclusions. They should be able to explain the author's use of literary devices such as irony or symbolism. When reading informative text, they should be able to apply text information appropriately to specific situations and integrate their background information with ideas in the text to draw and support conclusions. When reading practical texts, they should be able to apply information or directions appropriately. They should be able to use personal experiences to evaluate the usefulness of text information.

Advanced 348

Twelfth grade students performing at the advanced level should be able to describe more abstract themes and ideas in the overall text. When reading text appropriate to 12th grade, they should be able to analyze both the meaning and the form of the text and explicitly support their analyses with specific examples from the text. They should be able to extend the information from the text by relating it to their experiences and to the world. Their responses should be thorough, thoughtful, and extensive. For example, when reading literary text, advanced-level 12th graders should be able to produce complex, abstract summaries and theme statements. They should be able to use cultural, historical, and personal information to develop and explain text perspectives and conclusions. They should be able to evaluate the text, applying knowledge gained from other texts. When reading informational text, they should be able to analyze, synthesize, and evaluate points of view. They should be able to identify the relationship between the author's stance and elements of the text. They should be able to apply text information to new situations and to the process of forming new responses to problems or issues. When reading practical text, advanced-level 12th graders should be able to make a critical evaluation of the usefulness of the text and apply directions from the text to new situations.

BEST COPY AVAILABLE

Mathematics

Figure 1.1 Description of Mathematics Achievement Levels for Basic, Proficient, and Advanced Fourth Graders

The five NAEP content areas are (1) numbers and operations, (2) measurement, (3) geometry, (4) data analysis, statistics, and probability, and (5) algebra and functions. At the fourth-grade level, algebra and functions are treated in informal and exploratory ways, often through the study of patterns. Skills are cumulative across levels—from Basic to Proficient to Advanced.

Basic 211

Fourth-grade students performing at the basic level should show some evidence of understanding the mathematical concepts and procedures in the five NAEP content areas. Fourth graders performing at the basic level should be able to estimate and use basic facts to perform simple computations with whole numbers; show some understanding of fractions and decimals; and solve some simple real-world problems in all NAEP content areas. Students at this level should be able to use—though not always accurately—four function calculators, rulers, and geometric shapes. Their written responses are often minimal and presented without supporting information.

Proficient 248

Fourth-grade students performing at the proficient level should consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content areas. Fourth graders performing at the proficient level should be able to use whole numbers to estimate, compute, and determine whether results are reasonable. They should have a conceptual understanding of fractions and decimals; be able to solve real-world problems in all NAEP content areas; and use four-function calculators, rulers, and geometric shapes appropriately. Students performing at the proficient level should employ problem-solving strategies such as identifying and using appropriate information. Their written solutions should be organized and presented both with supporting information and explanations of how they were achieved.

Advanced 280

Fourth-grade students performing at the advanced level should apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content areas. Fourth graders performing at the advanced level should be able to solve complex and nonroutine real-world problems in all NAEP content areas. They should display mastery in the use of four-function calculators, rulers, and geometric shapes. These students are expected to draw logical conclusions and justify answer and solution processes by explaining why, as well as how, they were achieved. They should go beyond the obvious in their interpretations and be able to communicate their thoughts clearly and concisely.

BEST COPY AVAILABLE

Figure 1.2 Description of Mathematics Achievement Levels for Basic, Proficient, and Advanced Eighth Graders

The five NAEP content areas are (1) numbers and operations, (2) measurement, (3) geometry, (4) data analysis, statistics, and probability, and (5) algebra and functions. Skills are cumulative across levels—from Basic to Proficient to Advanced.

Basic 256

Eighth-grade students performing at the basic level should exhibit evidence of conceptual and procedural understanding in the five NAEP content areas. This level of performance signifies an understanding of arithmetic operations—including estimation—on whole numbers, decimals, fractions, and percents. Eighth-graders performing at the basic level should complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs. They should be able to solve problems in all NAEP content areas through the appropriate selection and use of strategies and technological tools—including calculators, computers, and geometric shapes. Students at this level also should be able to use fundamental algebraic and informal geometric concepts in problem solving.

As they approach the proficient level, students at the basic level should be able to determine which of available data are necessary and sufficient for correct solutions and use them in problem solving. However, these 8th graders show limited skill in communicating mathematically.

Proficient 294

Eighth-grade students performing at the proficient level should apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas. Eighth-graders performing at the proficient level should be able to conjecture, defend their ideas, and give supporting examples. They should understand the connections between fractions, percents, decimals, and other mathematical topics such as algebra and functions. Students at this level are expected to have a thorough understanding of basic level arithmetic operations—an understanding sufficient for problem solving in practical situations.

Quantity and spatial relationships in problem solving and reasoning should be familiar to them, and they should be able to convey underlying reasoning skills beyond the level of arithmetic. They should be able to compare and contrast mathematical ideas and generate their own examples. These students should make inferences from data and graphs; apply properties of informal geometry; and accurately use the tools of technology. Students at this level should understand the process of gathering and organizing data and be able to calculate, evaluate, and communicate results within the domain of statistics and probability.

Advanced 331

Eighth-grade students performing at the advanced level should be able to reach beyond the recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles in the five NAEP content areas. Eighth-graders performing at the advanced level should be able to probe examples and counterexamples in order to shape generalizations from which they can develop models. Eighth-graders performing at the advanced level should use number sense and geometric awareness to consider the reasonableness of an answer. They are expected to use abstract thinking to create unique problem-solving techniques and explain the reasoning processes underlying their conclusions.

BEST COPY AVAILABLE

Figure 1.3 Description of Mathematics Achievement Levels for Basic, Proficient, and Advanced Twelfth Graders

The five NAEP content areas are (1) numbers and operations, (2) measurement, (3) geometry, (4) data analysis, statistics, and probability, and (5) algebra and functions. Skills are cumulative across levels—from Basic to Proficient to Advanced.

Basic 287

Twelfth-grade students performing at the basic level should demonstrate procedural and conceptual knowledge in solving problems in the five NAEP content areas. Twelfth-grade students performing at the basic level should be able to use estimation to verify solutions and determine the reasonableness of results as applied to real-world problems. They are expected to use algebraic and geometric reasoning strategies to solve problems. Twelfth-graders performing at the basic level should recognize relationships presented in verbal, algebraic, tabular, and graphical forms; and demonstrate knowledge of geometric relationships and corresponding measurement skills.

They should be able to apply statistical reasoning in the organization and display of data and in reading tables and graphs. They also should be able to generalize from patterns and examples in the areas of algebra, geometry, and statistics. At this level, they should use correct mathematical language and symbols to communicate mathematical relationships and reasoning processes; and use calculators appropriately to solve problems.

Proficient 334

Twelfth-grade students performing at the proficient level should consistently integrate mathematical concepts and procedures to the solutions of more complex problems in the five NAEP content areas. Twelfth-graders performing at the proficient level should demonstrate an understanding of algebraic, statistical, and geometric and spatial reasoning. They should be able to perform algebraic operations involving polynomials; justify geometric relationships; and judge and defend the reasonableness of answers as applied to real-world situations. These students should be able to analyze and interpret data in tabular and graphical form; understand and use elements of the function concept in symbolic, graphical, and tabular form; and make conjectures, defend ideas, and give supporting examples.

Advanced 366

Twelfth-grade students performing at the advanced level should consistently demonstrate the integration of procedural and conceptual knowledge and the synthesis of ideas in the five NAEP content areas. Twelfth-grade students performing at the advanced level should understand the function concept; and be able to compare and apply the numeric, algebraic, and graphical properties of functions. They should apply their knowledge of algebra, geometry, and statistics to solve problems in more advanced areas of continuous and discrete mathematics.

They should be able to formulate generalizations and create models through probing examples and counterexamples. They should be able to communicate their mathematical reasoning through the clear, concise, and correct use of mathematical symbolism and logical thinking.

BEST COPY AVAILABLE